

BAB VII. KESIMPULAN DAN SARAN

7.1. Kesimpulan

Berdasarkan hasil penelitian yang dilakukan beserta hasil dari berbagai pengujian didapatkan, dapat ditarik kesimpulan sebagai berikut:

1. Untuk mengolah data kotor menjadi data kluster terbaik, dibutuhkan beberapa proses yaitu preprocessing untuk menyiapkan data sebelum diimplementasikan kepada algoritma, vectorizing menggunakan TF-IDF untuk mengubah data menjadi vektor, clustering menggunakan Kmeans untuk mengelompokkan data, dan testing menggunakan silhouette coefficient untuk menemukan data jumlah kluster yang terbaik.
2. Menggunakan TF-IDF dan Kmeans dalam membantu dosen mengelompokkan jawaban uraian mahasiswa ternyata cukup efektif karena dengan persentase kesesuaian hasil pengelompokkan sebesar 65% dosen dapat mengelompokkan jawaban uraian dengan waktu jauh lebih cepat dari mengelompokkan jawaban uraian secara manual.
3. Kualitas penggunaan TF-IDF dan Kmeans dalam mengelompokkan jawaban uraian mahasiswa masih cukup rendah, dengan didapati nilai silhouette coefficient terbesar sebesar 0.12 yang menunjukkan bahwa metode yang dipakai untuk mengolah data jawaban uraian mahasiswa bukan merupakan metode yang paling efektif.

7.2. Saran

Saran yang dapat diberikan dari penelitian ini untuk penelitian selanjutnya adalah sebagai berikut:

1. Pada *Kmeans clustering* dan *silhouette coefficient* dapat dilakukan perhitungan jarak menggunakan metode lain seperti *cosine similarity*.
2. Dapat dilakukan pengujian kluster menggunakan metode lain seperti *inertia*, *dunn index*, *davies-bouldin index*, dan *calinski-harabasz index*.
3. Sistem dapat ditambahkan fitur untuk mengelompokkan jawaban uraian yang mengandung gambar.